Table of Contents

Protocols to measure *creation* of knowledge and *contributions* to its value, in a trusted, protected, and transparent manner.

Knowledge Protocols Overview

- **PROTOCOLS** to describe quantification of the knowledge in content and distributing the rights to the content in any format viz; text, media, statistical data sets, etc.
- **KNOW: K**nowledge and **N**otion **O**f its **W**orth: Evaluating and quantifying the Knowledge (Value) stored in any DATA / Information format.
- **FORK**: **F**ractional **O**wnership and **R**ights to **K**nowledge: A Distributed Authoring and Versioning of knowledge using blockchain technology.

Knowledge protocols are defined to quantify 'value of information' and provide mechanisms to 'attribute its ownership' to authors using blockchain technology. It enables transparent collaborative content authority, distributed ownership and incentivization opportunities for creators viz; educators, researchers, artists, marketeers, journalists and even organizations; both academic and business.

As a Technology (Utility Token) it can enable use-cases in education, corporate L&D, training, news and journalism, marketing, research, science and more. It is purposefully built with the aim of decentralization for transparent assignment and share of ownership between collaborators.

1. The data economy in the age of information

THE INFORMATION AGE

THE VALUE OF KNOWLEDGE IN A DATA ECONOMY

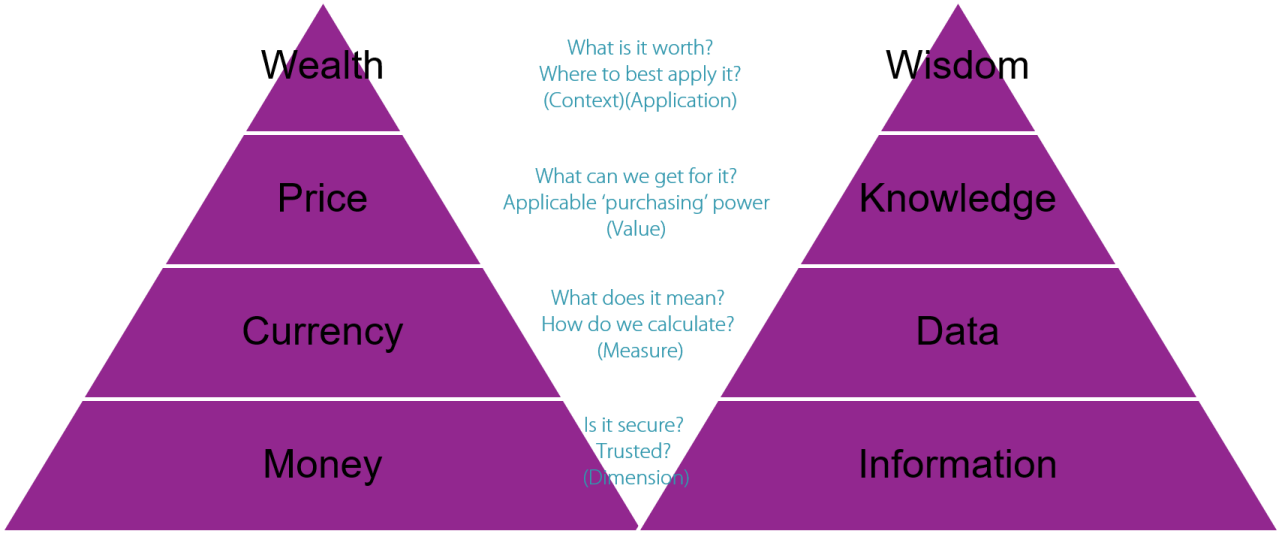
Not so long ago in history, Gold was the globally accepted store of value when the British established a mint which was ascertained to produce 1 ounce of gold repeatedly with precision and accuracy. Later, it was naturally opted as the reserve pegged to the paper currency. In the measure of gold, this currency was printed, and wealth was calculated. The need for minting the currency soon outpaced the actual rate of mining of gold. Hence the peg was naturally broken, and money was free to rein as long, and as abundant, as paper and ink. The currency was now artificially controlled and depleting in value at a faster rate and the 'proof of work' of mining gold transformed into 'proof of stake' by the trust placed in the local financial system rather than the currency itself. The scale of measurement changed, and value stored in each unit of currency is just about an abstract understanding of its purchasing power. If you are reading a blockchain whitepaper then chances are you already know the story by heart and in depth.

Now let us contextually change everything in the analogy in Data.

In the age of information,

DATA is the CURRENCY that is exchanged (TRANSACTED) to obtain VALUE.

KNOWLEDGE is the VALUE of DATA



THE ABUNDANCE:

Both data and currencies are being minted on an *exponential pace*.

90% of $$ ever printed in history is printed in the last year. (Source)

In all Human history, the information created till the information age, is about X ZB, Data generated in 2010 reached 1 ZB of digital data, by 2018 it was 33ZB and by 2025 it is expected to grow to 175 zettabytes by 2025, according to a recent IDC and Seagate report. Leveraging the power of this data, the next era of virtual and intelligent augmentation is not linear but exponentially huge.

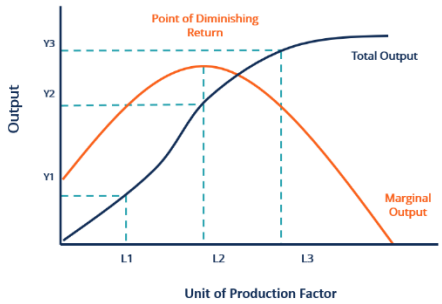
Considering these statistics (source CISCO VNI: Forecast and Methodology, 2015-2020). Global IP traffic will increase nearly threefold over the next 5 years, **annual global IP traffic will reach 2.3 ZB per year by 2020**, **smartphone traffic will exceed PC traffic by 2020**, and **the number of devices connected to IP networks will be more than three times the global population by 2020**. While it's undeniable that **the data volumes are growing faster than ever before** and we will continue to create new content to be managed every second.

Yet according to McKinsey (Cited) less than 1% of the data is analysed.

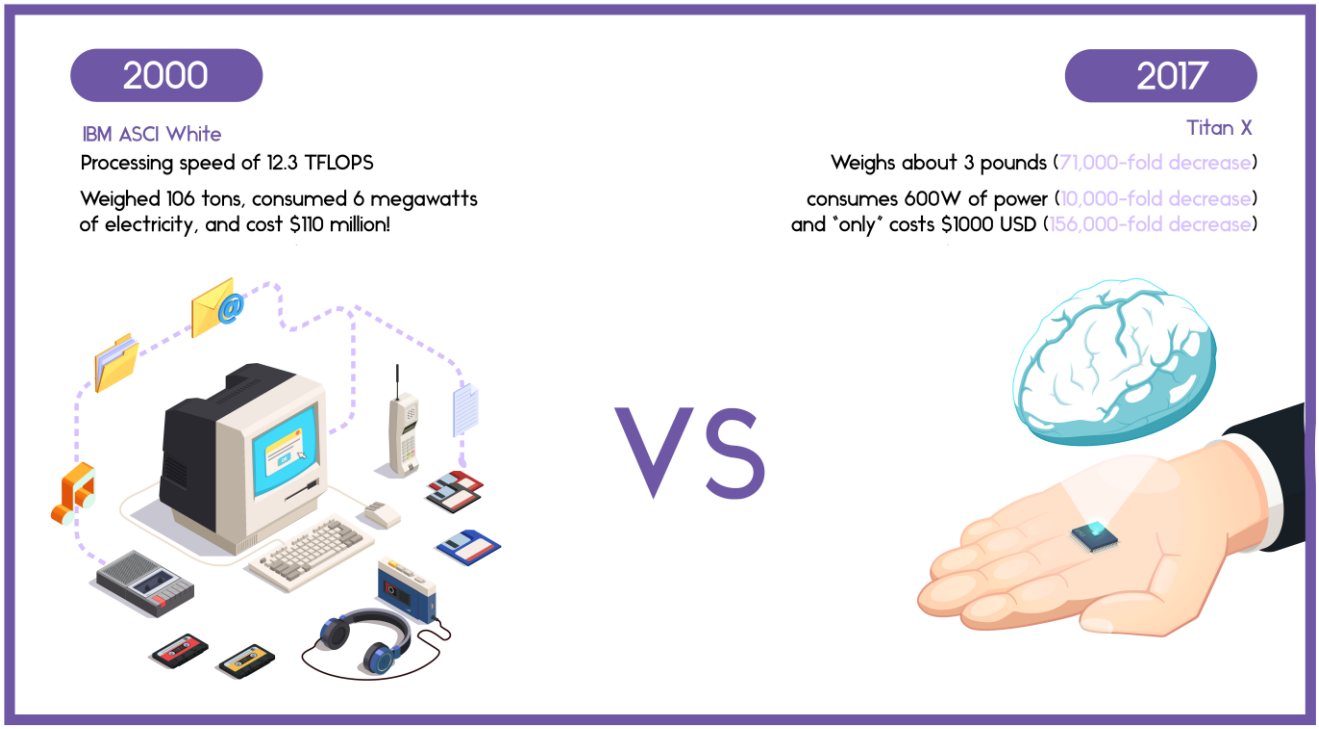
THE LAW OF DIMINISHING RETURNS:

Data is increasing day by day and there is a lot of noise in data. As same as money, you can increase the data set size however, the 'value' the data stores in it, as in the 'information' that can be drawn out, will not change as it is dependant not on the 'size', but the 'type' of information.

For example, what you want to derive from a 1000 people data set as % of people getting married and divorced will not deviate much even if you increase the data set to 10000 or 1000000 in the same local environment. So, one mayt easily argue that the returns diminish while increasing data. The knowledge stored in the data will remain unchanged within a margin. Deploying data driven decisions place organizations in with a strategic edge yet leading data analytics organizations are learning the reality that the value of data peaks and then plateaus against the volume of data.



To compare to the growth in data processing capabilities, consider the fastest supercomputer on Earth in the year 2000, the IBM ASCI White, which had a theoretical processing speed of 12.3 TFLOPS. That computer weighed 106 tons, consumed 6 megawatts of electricity, and cost $110 million ($156 million in 2017 USD), whereas the Titan X weights about 3 pounds (a 71,000-fold decrease), only consumes 600W of power (a 10,000-fold decrease), and "only" costs $1000 USD (a 156,000-fold decrease). Fitting the computing power of a 2000-era supercomputer on a single chip in the year 2017 is an incredible change and there is no sign of the pace of improvement slowing down any time soon. In other words, data gets cheaper and cheaper to collect, to process, and the more we collect, as its value gets distributed.



In perspective of knowledge, the more abundant the knowledge the more is its applicable value, 'wisdom' but cheaper to acquire. Extracting new knowledge out of information increases its value, and sharing it decreases the per unit value corresponding to its spread. Say if everyone on the planet knows something then the 'Knowledge' as a whole is very valuable, yet that value is now distributed across 7 billion people who all know it so it's very cheap to acquire it or use.

Galileo discovering the knowledge of Jupiter's biggest moons would be wisdom, yet every kid who can name all 4 of them now would not be considered as wise. So, if everyone knows about the solar system and the planets and bodies within it, the information is abundant enough that it is of lesser contextual value and is basic knowledge that can be used cheaply to educate at basic elementary level. Yet, knowing how much water is on Europa or what is the soil composition of Mars crust is of good measurable value to plan space missions of the near future. New information keeps coming in at an exponential rate as newer and newer discoveries keep adding to it. Thus, knowledge has a way of ***'decreasing in value over time with its spread'*** and has a way of ***'building on top of it as way of adding information'***.

THE REAL 'VALUE'

So how do we measure the added value and the depleted value of knowledge to a system? The problem is familiar to anyone who has worked with concurrent or distributed computing. Think about GIT. Many developers working on a single codebase. We have some shared state (the code base) with developers making updates in parallel. We need to somehow combine these by serializing the changes into some consensus update (or rollout). Our task is made more complicated by the fact that getting a system to execute and run correctly implies very complex validity criteria for that shared state. There's no way of creating a deterministic algorithm to find consensus. Humans need to find the consensus, and that consensus may involve mixing choice parts of different updates. Often consensus can only be reached with original updates to resolve the conflicts.

Value exists by Consensus and confirmed only when it is realized in a transaction

Let's move on from developer code and think of knowledge in general. The 'knowledge' (Value) stored as 'Information' in 'Data'.

The problem of calculating Value in Data can be distilled into four primary categories: **time**, **cost**, **quality**, and **complexity**. Each aspect increases as the data increases. One certainly affects others. Often it is the case that solving for only one aspect increases the rest. Therefore, the answer lies in addressing all in a holistic manner.

The VALUE of the Information is corelated to the following metrics.

1. Complexity to learn or acquire (Difficulty to learn or extract knowledge)
2. Complexity to share (Difficulty to teach or impart knowledge)
3. Complexity to apply (Benefit from Knowledge)
4. Demand vs Abundance (Scarcity)
5. Time needed to process (acquire or learn from the knowledge)
6. Recency and latency

There are many subjective arguments that can be made to above metrics also. The commercial applicability isn't (and shouldn't be) the only factor of calculating wisdom.

Time it takes to learn the subject to be able to demonstrate application. Pure Mathematics is inapplicable?

La . Lb

https://www.grin.com/document/501331#:~:text=Catalog%20%3E%20Philosophy%20%3E%20Miscellaneous-,%22The%20quality%20of%20knowledge%20is%20best%20measured%20by%20how%20many,knowledge%20and%20real%20life%20examples.

Agreeable Value of Knowledge:

Knowledge is considered as an intangible asset, something that is not quantifiable yet with a single unit or a single applicable formula. Yet we can agree that it is traded between people with an agreed **consensus** and is confirmed by some form of a **transaction**. It could be consensus and transaction between a teacher and a student, in hardware in the form of embedded knowledge within technological systems, or even between programs communicating in a defined manner exchanging knowledge in form of AI and machine learning.

THE INFORMATION OWNERSHIP

THE OWNERSHIP OF KNOWLEDGE IN A DATA ECONOMY (IP)

KNOWLEDGE protocol establishes a manner to contextually quantify the value locked in any form of data and to compliment, FORK is written to assign, protect and distribute the rights to its owners. Together, they form the foundations for **contextual applications** that *measures* the *value of information* as '**Knowledge'** and securely *assigns ownership* to '**information**'.

FORK Protocol will protect the rights of ownership and 'contributorship' of the information by creating content IDs and writing it on a decentralized ledger with blockchain technology.

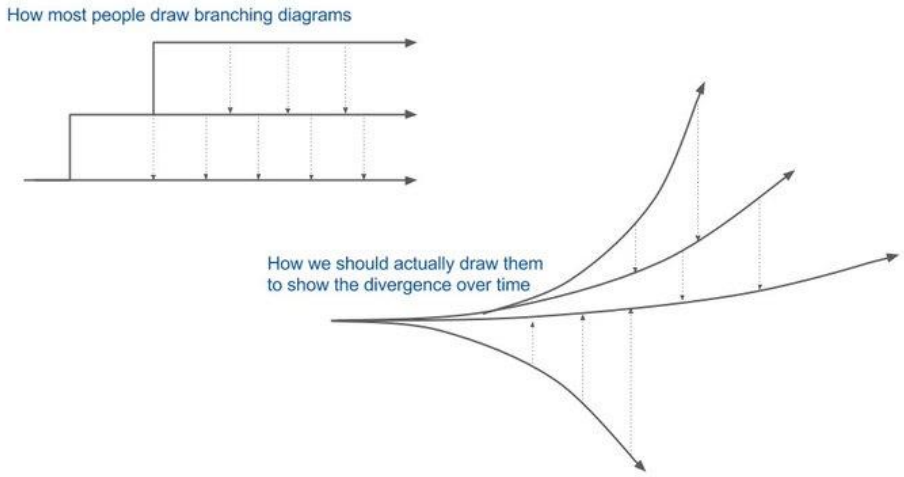
FORK Smart Contract	KNOWLEDGE Smart Contract
FORK is an IP protection protocol written for distribution of content ownership. Written on chain as part of FORK protocol record.	KNOWLEDGE is a utility token that commercializes the value stored in the content. Content Value is measured in 'KNOW' and is written on chain in KNOWLKEDGE LEDGER.
FORK allows rights of ownership to its contributors in an agreed percentage distribution. One (content ID) to many (Owners) relationship	KNOWLEDGE is the quantified value of the creation or contribution a creator has generated in the ecosystem. One (Content ID) to One (Knowledge value) to relationship.
Focused on protection and distribution of rights to knowledge	Focused on valuing/monetizing the content and rewarding the creator
Maintains verifiable records and version-controlled distribution of share of ownership with each revision of content on-chain	Maintaining privacy and security for transactions on-chain.
Functions to record content modification history, IP assignments, ownership distribution, ~~value~~, etc.	Functions to manage transactions; like direct transfers, sale, and purchase of content in marketplace, staking, rewards, etc.
Security focus on counterfeit, fake content, copyright claims, piracy, impersonation, etc.	Security focus on privacy, secure transactions, wallet attacks, phishing, etc.
Risks: Viewability problems/attribution/Identity theft	Risks: Secure attribution/monetary theft
Smart Contract -	Smart Contract -

Struct - Content DID, Version ID, Meta Data, Owners (Vector (Expandable Array with Owner DID and Share in %), Revised Knowledge Value (Integer)) FORK contract address, # Transaction Address	Struct - Content DID, Genesis V1 # (Access Key), Meta Data, Value KNOW contract address, #transaction address

Blockchain technology allows verifiable means of ownership to digital assets. Ownership of Bitcoin, as the first cryptocurrency, solved this problem of uniquely owning a digital asset. In the Bitcoin system, you "control" an "address" (or wallet address) where your crypto is stored. You hold the "private key" to that address. A private key is like a password – that is for you to 'own' and 'access' your crypto. Anyone with the key is the owner, there is no mechanism to reclaim this or verify by any other identity mechanism. There's a single shared global database that keeps track of who owns what, and can then easily prevent people from spending the same token twice.

"Ownership" is having rights to knowledge like as in an IP, Patent, or protected copyrights. "Owning" knowledge means you hold the private key to your information and can control how to share it. This gives you the credible right to transfer that knowledge to others. With Knowledge protocol you can author information (create), and with FORK

protocol you can claim ownership. The way FORK (Fractional Ownership and Rights to Knowledge) is managed is that you can share the ownership with other contributors and add value to others work by forking it (in the git-flow context).



Jonny LeRoy likes to point out this flaw in how people draw branching diagrams.

INFORMATION SHARING

THE TRANSACTIONS OF KNOWLEDGE IN A DATA ECONOMY

In the age of INFORMATION, DATA is the CURRENCY that is exchanged (TRANSACTED) to obtain VALUE. Money is any trusted store of VALUE that helps us make "TRADE", wherever we need to TRANSACT VALUE of goods and services.

• Currency is a measure of MONEY that is LOCALLY accepted • Can be borrowed, traded, staked, or lent at interest. • Is subjected to inflation, can be staked to grow, or earn capital interest by investing and applying to businesses.	• Data is a measure of INFORMATION that is CONTEXTUALLY applicable • Can be sourced, exchanged, cited, or further developed. • Is subjected to inflation by losing value over time, can be improved, distilled, and applied.

After defining the protocol of usage, it is helpful to consider what types of data you create and/or work with, and what format those data take. Your data stewardship practices will be dictated by the types of data that you work with, and what format they are in.

Data Types

Data types generally fall into five categories:

Observational
- Captured *in situ*
- Cannot be recaptured, recreated, or replaced
- Examples: Sensor readings, sensory (human) observations, survey results

Experimental
- Data collected under controlled conditions, *in situ* or laboratory-based
- Should be reproducible, but can be expensive
- Examples: gene sequences, chromatograms, spectroscopy, microscopy

Derived or compiled
- Reproducible, but can be very expensive
- Examples: text and data mining, derived variables, compiled database, 3D models

Simulation
- Results from using a model to study the behavior and performance of an actual or theoretical system
- Models and metadata, where the input can be more important than output data
- Examples: climate models, economic models, biogeochemical models

Reference or canonical
- Static or organic collection [peer-reviewed] datasets, most probably published and/or curated.
- Examples: gene sequence databanks, chemical structures, census data, spatial data portals.

Data comes in various formats: text, numeric, multimedia, models, software languages, discipline specific (e.g. crystallographic information file (CIF) in chemistry), and instrument specific. (source)

TYPE OF DATA	PREFERRED FILE FORMATS FOR SHARING, RE-USE AND PRESERVATION	Other Acceptable formats
Quantitative tabular data with extensive metadata • **a dataset with variable labels, code labels, and defined missing values, in addition to the matrix of data**	• SPSS portable format (.por) • delimited text and command ('setup') file • (SPSS, Stata, SAS, etc.) containing metadata information • structured text or mark-up file containing metadata information, e.g. DDI XML file	MS Access (.mdb/.accdb)
Quantitative tabular data with minimal metadata • **a matrix of data with or without column headings or variable names, but no other metadata or labelling**	• comma-separated values (CSV) file (.csv) • tab-delimited file (.tab) • including delimited text of given character set with SQL data definition statements where appropriate	• delimited text of given character set -- only characters not present in the data should be used as delimiters (.txt) • widely-used formats, e.g. MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf) and OpenDocument Spreadsheet (.ods)
Geospatial data vector and raster data	• ESRI Shapefile • (essential: .shp, .shx, .dbf ; optional: .prj, .sbx, .sbn) • geo-referenced TIFF (.tif, .tfw) • CAD data (.dwg) • tabular GIS attribute data	• ESRI Geodatabase format (.mdb) • MapInfo Interchange Format (.mif) for vector data
Qualitative data textual	• eXtensible Mark-up Language (XML) text according to an appropriate Document Type Definition (DTD) or schema (.xml) • Rich Text Format (.rtf) • plain text data, UTF-8 (Unicode; .txt)	• plain text data, ASCII (.txt) • Hypertext Mark-up Language (HTML) (.html) • widely-used proprietary formats, e.g. MS Word (.doc/.docx) • LaTeX (.tex)
Digital image data	TIFF version 6 uncompressed (.tif)	• JPEG (.jpeg, .jpg) • TIFF (other versions; .tif, .tiff) • JPEG 2000 (.jp2) • Adobe Portable Document Format (PDF/A, PDF) (.pdf)
Digital audio data	• Free Lossless Audio Codec (FLAC) (.flac) • Waveform Audio Format (WAV) (.wav) • MPEG-1 Audio Layer 3 (.mp3) - spoken word audio only	• MPEG-1 Audio Layer 3 (.mp3) • Audio Interchange File Format (AIFF) (.aif)
Digital video data	• MPEG-4 High Profile (.mp4) • motion JPEG 2000 (.jp2)	JPEG 2000 (.mj2)
Documentation & Scripts	• Rich Text Format (.rtf) • Open Document Text (.odt) • HTML (.htm, .html)	• plain text (.txt) • widely-used proprietary formats, e.g. MS Word (.doc/.docx) or MS Excel (.xls/.xlsx) • XML marked-up text (.xml) according to an appropriate DTD or schema, e.g. XHMTL 1.0 • PDF/A or PDF (.pdf)
Chemistry data: spectroscopy data and other plots which require the capability of representing contours as well as peak position and intensity	Convert NMR, IR, Raman, UV and Mass Spectrometry files to JCAMP format for ease in sharing. JCAMP file viewers: JSpecView, ChemDoodle	

Formats more likely to be accessible in the future are:
- Non-proprietary
- Open, documented standards
- In common usage by the research community
- Using standard character encodings (ASCII, UTF-8)
- Uncompressed (desirable, space permitting)

Knowledge in the proper context can be immensely beneficial yet it mostly remains locked up in data. There is a wide gap between data collected and data processed and remains uncontextualized due to below reasons:

- Difference of costs between data storage (low) and data processing (high)
- Data silos
- Data privacy and ill intent usage
- Absence of a traceability mechanism
- Single point ownership mechanisms that do not allow sharing or contributions to data
- Lack of commercialization standards that can help define value locked in data or value added to data
- Lack of data trails or version control to audit data quality, licencing, rights, and usage
- Control and regulations that are loosely defined and mostly misinterpreted that miss the opportunity to open data
- Defined and trusted scale of pricing data

Once data is shared, it is out of the hands of owner and in the wild. Without measures that can provide a semblance of control, an audit trail on usage, and fair compensation schemes, the data will remain locked up.

- *Our friends at Ocean*

(Re)building trust and transparency in information (Knowledge) sharing.

It is well established fact now that people, governments, and organizations, alike are hesitant to share data as more and more cases of data abuse surface showing exploitation of user trust and confidentiality.

Most of the times that you see the pop-up of 'We Value your Privacy', it is basically asking your consent to violate it. Publishers and content creators are shutting down or retaliating with self-destructive tactics as users enable mechanisms like VPNs, ad-blockers, fake impersonations in response to privacy violations, personalization, and targeted advertisements. Fake News, bad content and Ad fraud are unchecked and unregulated medium of reaching out to masses. Educators, publishers, advertisers struggling to find real solutions that comply with intent of ePrivacy/GDPR regulations. This is a fundamentally unsustainable future for the next web generation.

What does not constitute Knowledge?

- (A) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and
- (B) Personnel and personally identifiable medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study."
- (C) "Digitally Recorded" material excludes physical objects (e.g., laboratory samples).

Take references from Wikimedia Unlock

2. Need (Problem-Solution fit)

To complete the comparison between the two pyramids and compare the information value to the monetary one we can see how wealth is defined in the real world and if information draws the same parallels. There are several mechanisms to evaluate wealth as asset value of a person or an organization.

- An asset is a thing that can be owned and controlled; and
- An asset is something that is exchangeable for cash; and
- An asset is something that can be transferred/shared/lent to generate probable future economic benefits.

Even though data/information/knowledge is not taken as an asset on the books <ref footnote>, it clearly has value as it meets all the above requirements and works in similar manner to how assets are valued.

wealthy and wise – Write something fancy.

<Start footnote> Most people recognize that information has value. The value is not always clear, but it is clearly there. Yet according to some professional organizations, digital information literally has no value. It sounds inexplicable, but it is true: Information has no official value in the eyes of the property and casualty (P&C) insurance industry. Nor does it have any official value to the accounting folk who set the Financial Accounting Standards (FAS).

Some of the companies in the Twin Towers who lamented not only the tragic loss of life and physical property also lost their data, so naturally what they did when they suffered some kind of loss was, they reached out to insurance companies, and they filed claims for the value of the information that they lost in the attacks."

However, the insurance companies denied the claims, saying they did not believe that information constitutes property and therefore were not covered under P&C policies. What is more, the insurance companies went so far as to update the commercial general liability (CBL) policy standard that's used by most insurance companies around the world to explicitly exclude information from P&C polices.

"When did they do that? - A month after 9/11."

Some of the courts have said information can be represented by bubbles on an optical drive or it can be printed and so forth, so it should be considered property. Other courts have said ridiculous things like, well electrons have negligible mass so information should not be considered property. - Excerpts from Laney's keynote at the Veritas Vision 2016 conference Las Vegas, Nevada Source <Close footnote>

Content as digital assets are electronically sold and purchased, you subscribe to digital information and services, and own information be it created as an IP or purchased as rights to a patent.

In today's world, content is of no shortage. Content has often been touted as the king. With everyone having the tools create content and the platforms to share it. Digital information has transformed itself into many kinds, viz; video, participative, immersive as 3D/AR/VR and gamification. Even though managing content in the digital age is a challenge with many dimensions, we have identified the main opportunities that offer people a collaborative knowledge creation and sharing experience. Even at a personal level oriented towards learning and self-growth while collaborating and connecting with information around our interests, passions, goals, and lifestyle.

We need a trusted mechanism to:

- Be able to define ownership and control our knowledge assets
- Be able to exchange it for cash
- Be able to transfer/share/loan to generate probable future economic benefits

Additional elements to provide safe environment/ecosystem to manage knowledge would have following attributes

Democracy - Each participant of the ecosystem has an equal and equitable access to creation, curation, distribution, consumption, and commercialization of knowledge (in form of content, data or learning material). Each participant of the community is incentivized of their contribution and participates in quality of the generated knowledge.

User Privacy - A vision of being a "trust-less" platform. You do not have to place your trust in any centralized organization to have possibility of being betrayed at any later stage. Authenticity is required for ownership however it is not built to be profiled or targeted in any ways in which the data is shared across the board.

Free Speech - A fundamental principle that should be protected online with disruptive tools.

Open Knowledge – Knowledge is not owned centrally by any entity but assigned back individually by the protocol to the creators and contributors in an equitable and decentralized manner in a blockchain ledger.

Traceability – Where the knowledge flows, who builds on it and how it is contextually used.

IP protected – Creators and contributors own the rights to knowledge they created. A decentralized ledger credits the contribution while distributing value amongst the creators for generated IP and transparent commercialization of that IP.

Customizability – Framework that others can tailor and build upon own social experiences to their liking.

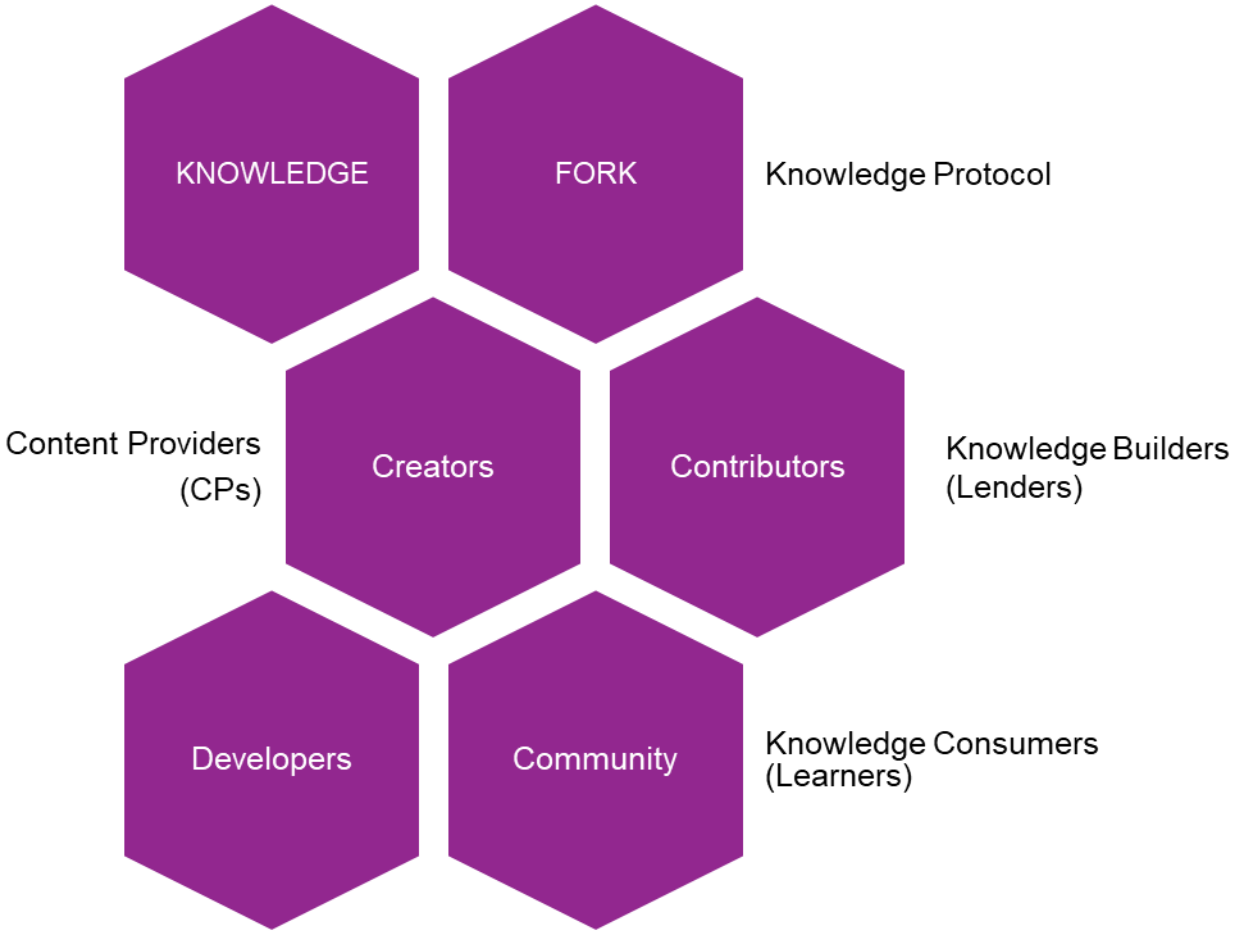
Information providers and custodians need to feel safe before they can share knowledge in the ecosystem. Apps built in the Knowledge network using the guidelines set forth in the protocol allow all actors to come forward and participate in the data economy freely, fairly and equitably.

---

3. Ecosystem

3.1 Operators and Participants

Network, Protocol, Applications and Stakeholders



Knowledge network is an ecosystem comprising of

- CODE: Technical guidelines as Knowledge and FORK protocols
- CONTENT: Knowledge builders as content providers, content creators and contributors
- COMMUNITY: The learners, developers, and governance framework

The knowledge network enables a safe and trusted ecosystem to share information while ensuring IP rights, guaranteeing version control, traceability for audits, and transparency while transacting in it.

The ecosystem described is a unique mix of technology that enables others to come and build solutions in top of it. It is less about what the tech "**does**", more of what the tech "**enables**".

Objectives of **CODE**:

Establish an automated and trusted mechanism of evaluating quality and value of contribution

The value scoring mechanism will be able to identify the value added by each contributor to one asset through versions and thus promote quality contributions and reward users accordingly. As the contributions of users are evaluated over their successive contributions, they will accumulate a score that reflects their reputation. This will allow a user's domain-specific expertise to be reflected within their contributions. The reputation will promote authenticity while mitigating the illicit social behaviour and continually improve the content quality. However, 'reputation' is different than 'influence', each content will still have to gain its authority organically. Just as a Wikipedia article is not affected in popularity of who authored it but by the community that cites it, approves it and reads it.

Objectives of **CONTENT**

At its core, the protocol derives its value from the ability to collaboratively create and own quality content on the knowledge network. FORK Protocol will protect the rights of ownership and contributorship of the informational content by creating content IDs and writing it on a blockchain decentralized ledger.

Content creators should be rewarded fairly for their knowledge and time through the original content they add to the platform. Creators will be able to tip one another for posts or comments and even receive pay-outs in KNOWLEDGE from the network for content that created quantifiable value. By building a trust-less transactional ecosystem enabling educators/creators to own authority and rights of rewards in this manner, we believe we will foster a "value" ecosystem. It enables equitable and equal revenue source for all.

Objectives of **COMMUNITY**

New applications will be able to integrate with FORK and KNOWLEDGE easily through the API interfaces allowing developers to integrate their applications to utilize the power of technology, the protocol, and the community.

4. Technology

4.1 Blockchain Solution

- Smart Contracts & Workflows

Creation - Knowledge focuses on “create/access” while FORK focuses on “read/edit” access permissions. But further variants could be: Unix-style (read, write, execute; for individual, group, all); database-style (CRED: create, read, edit, delete), or blockchain database-style (CRAB: create, read, append, burn). When Knowledge is created, ie; a dataset identified and brought to the system an initial version is written with Meta Data

Struct - Content DID, Genesis V1, Meta Data, Total Value, Complexity Indicator, Scarcity Indicator,
This information is written on chain using KNOWLEDGE smart contract, #Transaction Address

Struct - Content DID, Version ID, Meta Data, Owners (Vector (Expandable Array with Owner DID and Share in %), Knowledge Value Integer)
FORK contract address, # Transaction Address

Contribution - We propose and describe our solution which utilizes Ethereum blockchain and smart contracts to approve, track, and carry out the version control functions for the content stored on IPFS. Our solution eliminates the need for a trusted centralized authority and provides transactions and records to share and keep track of different versions of online information (data, documents, and media) with high integrity, resiliency, and security.

Access – Using Data access tokens by Ocean protocol. Data access is always treated as a data service. This could be a service to access a static dataset (e.g. a single file), a dynamic dataset (stream), or for a compute service (e.g. “bring compute to the data”). For static data, we can tune variants based on the type of storage: Web2 cloud (e.g. AWS S3), Web3 non-permanent (e.g. Filecoin), Web3 permanent small-scale (e.g. Ethereum), Web3 permanent large-scale (e.g. Arweave), or go meta using IPFS but “pinned” (served up) by many places. For dynamic data, variants include Web2 streaming APIs (single-source), Web3 public data oracles (e.g. Chainlink), and Web3 private data oracles (e.g. DECO).

- Intended Product Architecture

- Decentralized Version Control of Information

The proposed solution tracks the state of the entity across its lifetime on the VCS. For a better explanation of the working of the proposed VCS, we will be using an example of a Document from which VideoWiki can be assumed to create Video Scenes. From here on out, asset 'X' shall refer to the said document and the proposed VCS shall track changes to its state as described below.

Also, while the functioning of the VCS is mentioned below, we need to consider the following two options of storage to be offered:

1. Use Storage on VideoWiki Side (like GitHub does).
2. Leave storage responsibility with the User (like Ocean does).

The effect of the choice of each of these options is explained further in the Pain Points Section. The storage option notwithstanding, the VCS is supposed to work to make sure there is transparency in the system and at the same time, there is RBAC (role-based access control) enforced (via smart contracts) in case privacy needs to be preserved (eg., private fork).

Content Creation and Update

When the user *Alice* uploads doc X for VideoWiki, the metadata is recorded.

Among other things, any transaction consists of the following:

1. **Hash of the content.** The hash is used to track the state changes of the X in its lifetime. Let the initial hash of X be **A.**
2. **Owner details.** This is the information about the user who owns data. This data can be the email or similar conventional IDs or a Decentralized Identifier.
3. **Flags**. Flags are fields in a transaction used to denote the metadata of the transaction. The flags used here are the following:
    - **Rollback: Possible values 0 and 1. If 1, then the current transaction is a record of a state change induced by rollback action. If 0 then the current transaction is just recording a new state change.**
    - **Steps: Possible values include all numbers in a set of Natural Numbers. This flag is checked iff Rollback Flag is 1. This denotes the number of steps to roll back to arrive at the state as denoted by the Hash field.**
    - **Merkle: Merkle tree hash to validate the rollback. This field is also checked iff Rollback is 1. This flag is discussed in the following sections.**

The hash can be created using any standard NIST approved hashing function like SHA256, MDA, and so on. Now, change to X can be in the following ways:

1. Owner *Alice* tries to modify X and the change is recorded as a new transaction that contains the new hash of X.
2. Another user *Bob* wants to contribute to X and the changes to X are recorded as a new transaction that contains the new hash of X.

Since the discussion of how another user may contribute to the asset has not been discussed formally, it has been left out of this draft at this time. However, we may follow Ocean's workflow where a potential buyer sends 1 data token to the owner of the data to indicate his/her interest in buying. In our VCS, we can implement a similar flow using our native token. More discussion on this is needed.

Rollback

At this point, we can assume that the original state of X has changed, and the new state is denoted by hash **B**. This basically completes once workflow, where a user adds an asset and its state, is tracked via its hash.

After multiple such successive commits, let the latest stage be **E** and the overall changes of X recorded in the system be:

A -> B -> C -> D -> E

This presents an opportunity to explore the commit rollback workflow. To demonstrate it, let us say that *Bob,* who was the contributor to these changes, decides to roll back to state **B**.

While this action can be easily facilitated in a centralized database by just deleting the transactions **E, D** and **C**, the same on Blockchain is not possible because of immutability. Thus, instead of CRED operations, we need to follow

CRAB operations (*Create, Read, Append* and *Burn*). To denote a rollback to a certain previous commit in a sequence of commits recorded, the resulting transaction, among other fields in its fields, has the following values in mentioned fields:

1. Hash: **B**
2. Owner: Bob's identity element.
3. Rollback: 1
4. Steps: 3
5. Merkle: f(C + f(**D** + f(**E**))), where f(x) is the hashing function used and the operator '+' denotes a simple text concatenation of hashes.

Thus, the effective state as tracked by the system is given by:

A -> B -> C -> D -> E -> <u>B</u>

Merge Flows

Merge Flows will need to account for scenarios where a contributor who has branched off the main asset wants to merge with either:

1. Same asset.
2. Different branched off assets.

In any case, two branches can only be merged at their latest stage. So piggybacking on the last example, if the asset X has 2 branches - one belonging to the creator of the asset while the other to a contributor and the creator has also made changes to his version/branch/state of the asset, then the merge will join together the latest branch stages.

The branch merge request will be initiated by either of the interested party. The request and the response will be recorded as a transaction on-chain encrypted via a multi-Sig. This will mean that if there are n potential branches merging, there would be n signatories - one from each branch. This is Shamir's Secret Key sharing technique, and the signatories will come to an agreement where the key distribution mandates a specific number of keys to come together to decrypt the document recorded on Blockchain.

A potential opportunity here could be allowing the provider (us) to hold a certain percentage of keys in-limbo. These would be termed as backup keys and may be used when there are disputes within the consortium of signatories or there is a loss of keys. Both flows are explored in later sections.

The merge will require an on-chain voting mechanism. This would take place in every signatory coming to a consensus about the share/ownership of the asset which would emerge from the merger of the two branches which may at the time of merger present two different products originating from the same source. A simple majority voting for deciding the credit/ownership/stake has been proposed. Every signatory would have an equal weightage in their vote. Special scenarios include the merger of two branches and the merger of multiple branches resulting in a tie. In scenarios involving only 2 signatories, each shall have the same weightage. However, in the case of a tie of n signatories (n>2), the owner of the official branch or origin, if present, reserves the kingmaker vote.

The asset emerging from this merger of branches would belong to an account whose keys are generated with the multi-sig approach. The keys themselves could be shared in a manner that reflects the stake/ownership over the asset as mentioned by the vote recorded on-chain. This would complete a merge flow resulting in the creation of a new asset.

<u>Deletion (Burn)</u>

The Right to Forget constitutes a very important fundamental right, especially in the European Union. This implies the deletion of all related data to specific information owned by the user and stored by the proprietor. However, this also collides with the basic tenet of Blockchain - immutability.

In such a case, it may be argued that nodes running our framework do not store the actual information but a cryptographic hash of it. If we were to take the example of an industry-standard hashing mechanism like SHA256, then the algorithm by design is a Merkle–Damgård construction. This means it is a one-way compression function. Given a hash, is not possible and infeasible to calculate the data which produced the hash. Thus, we arrive at 2 scenarios:

1. **The User stores the data.** If the user is responsible for maintaining their own data and the blockchain-based DVCS is utilized as a trusted state verification system then when the user invokes their Right to Forget (by requesting a delete), only marking their specific branches as being burned suffices. This can easily be done by adding in a commit at the end of the branch they requested deletion of. This commit would globally signify deletion in our whole system and may be used as a flag and restrict further branching.
2. **We store the data.** If we store the data, then we need to provide proof of Burn. This would signify that we relinquish our rights to access the data. Storing data on our site does provide us with a way to maintain a GitHub-flow structure. However, it also makes us liable to data lawsuits and binding to data regulatory compliance. Thus, this option might not be the best option when starting out.
3. **Zero Knowledge Proofs.** Data is stored with us but is confirmed and deleted with a zero-knowledge proof mechanism. We only share access (link) to the data and the key is burnt which essentially marks that the data is lost and is irrecoverable without its key to access. The access to data whether it exists or it is burnt can be verified publicly using zero knowledge proofs.

- Decentralized Storage Network (IPFS)
- Proof-of-Replication and Proof-of-Spacetime
- Smart Contracts
- Integration with Other Systems
- Proofs and Formal Verification

3.3 Token Generation Event/Token Economics

- Token Generation and Allocation

Why a new token?

No blockchain today has adequate performance for use as a full-scale distributed human review system, anyone familiar with the network fees on blockchains would agree to that. Ethereum has the most robust smart contract support of any popular blockchain, but without additional development it is too slow and expensive for many applications. Plasma, Lightning, Hashgraph, etc are still orders of magnitude away from necessary cost/speed performance and not yet robust. Future improvements may eventually make this feasible, but these are still early days for networks.

Knowledge is created and distributed in ‘Knowledge’ tokens.

Struct - Content DID, Genesis V1, Meta Data, Total Value, Complexity Indicator, Scarcity Indicator,
This information is written on chain using KNOWLEDGE smart contract, #Transaction Address

Struct - Content DID, Version1# ID (access key), Meta Data, Owners (Vector (Expandable Array with Owner DID and Share in %), Revised Knowledge Value (Integer))
FORK contract address, # Transaction Address

- <Token Name> as a Utility token

Our ‘Knowledge’ Token contract thus implements a custom batch API that supports efficient micro-assignments and micro-payments via one-to-many transfers. In other words, a single call can specify pay-outs from a smart bounty to 1000 addresses. This enables new and interesting use cases while remaining EIP20-compatible. We will open sourced

the audited contract with a library and comprehensive test suite to help other projects in the wider Ethereum developer community adopt this approach as it suits their needs.

What is Knowledge?

Knowledge is a protocol defined to replace the outdated and fragmented management of information systems with a quantified, source and version controlled ledger, governed by contributors.
The network can be hosted on any L1 blockchain with a native token of KNOWELDGE, which users earn by creating or contributing to knowledge. The token gets burnt on the distribution or transfer of knowledge thus mimicking the world system where knowledge created has value, and distribution of knowledge spreads that value amongst more people making it more and more common.

Knowledge is created and distributed in 'Knowledge' tokens.

> Struct - Content DID, Genesis V1, Meta Data, Total Value, Complexity Indicator, Scarcity Indicator,
> This information is written on chain using KNOWLEDGE smart contract, #Transaction Address
>
> Struct - Content DID, Version1# ID (access key), Meta Data, Owners (Vector (Expandable Array with Owner DID and Share in %), Revised Knowledge Value (Integer))
> FORK contract address, # Transaction Address

Our 'Knowledge' Token contract thus implements a custom batch API that supports efficient micro-assignments and micro-payments via one-to-many transfers. In other words, a single call can specify pay-outs from a smart bounty to 1000 addresses. This enables new and interesting use cases while remaining EIP20-compatible. We will open sourced the audited contract with a library and comprehensive test suite to help other projects in the wider Ethereum developer community adopt this approach as it suits their needs.

Use Cases:

Products that are involved in creation, management or distribution of knowledge would be able to use the knowledge protocols to quantify and commercialize the value in the information/data they manage.

Knowledge Creation Products that involve in creation of knowledge, like contributory research, video trainings and elearning, music ownership, etc. could use knowledge as a mechanism to assign value to the creation. Enabling exchange, sale and trade of content in the unit of knowledge.	Knowledge Distribution Products involved in training, coaching, learning courses and marketplaces, etc. could use knowledge tokens to facilitate payment for the streams of knowledge. Enabling access to live streams, classes, learning courses and educational material.	Knowledge Access Access to knowledge can be granted to the holders of knowledge token in form of source forks, programmed information or data sets. Enabling access to data without providing actual data or any transfer of ownership.	Zero Knowledge Proofs Managing knowledge proofs on the network without providing actual access to data or releasing control or ownership. Zero knowledge proofs have other utilities in the environments with AI data sets and programmed smart contracts.

The community, through apps like VideoWiki (creation of content) and GetBoarded (Distribution and L&D) have already reached to 12,000 subscribers and more than 30,000 participants.

Token details

> Ticker: KNOW
> Standard: ERC 20
> Max Supply: 1,000,000,000

3.4 Technology Roadmap

5. The Application

The first application to be using the knowledge protocols is Video.Wiki.

VideoWiki is an immersive encyclopedia, driven and sustained by a global community.

It is an Open Collaboratory Content Editing Platform that enables rapid Creation, Modification, Protection, and Monetization of Immersive Content.

- Assisted immersive content generation using AI without advanced tool or tech knowledge
- Collaborative Authoring of content with Distributed Ownership and Share of Financial Renumeration
- Reputation and Influence scores for the participants that allow for community intelligence for content ranking and moderation

A Web 3 app using knowledge protocols

The term, coined by the reporter John Markoff of The New York Times in 2006, refers to a new evolution of the web, its third generation, and includes specific innovations and practices. It's easier to identify the major differences between **Web 1.0** (users passively consult web pages and for the most part don't participate in generating content) and **Web 2.0** (users create content and interact with sites and with each other through social media, forums, etc.). Instead, with **Web 3.0**, the differences are not as clearly defined however below are the key characteristics of our applications.

Open: Open Source, crowd developed, and crowd supported content

Collaboratory: multiple people can come together to collaborate on each content being created from anywhere in the world

Content Editing: To create content that is fast, AI assisted and in an immersive media format enabling:

- **Rapid Creation**
- **Modification**
- **Authentic and Protected**
- **Monetizable**
- **Immersive and Engaging**

- **Protected**
- **Reusable**
- **Auditable**

Transactional (**Financially Rewarding**): A mechanism to credit quality authorship and discredit content piracy by creating a decentralized ledger to authenticate the creation and contributions to media format. The protocol enables a financial ecosystem of P2P monetary exchange between creator/contributors and consumers.

6. VideoWiki Solution and ‘Knowledge’ as a Platform Utility Token

4.1 Productization of the Technology

- Technology Objectives
- Market Potential
- Core Values and Impact of Technology
- Digital Transformation Advantage Matrix

4.2 Product Overview

- Beta Product
- TestNet

7. User Types

- Creators/~~Collaborators~~
- Contributors
- Curators
- Critics
- Coders
- Committee
- Consumers
- Companies

1. **Creator**: Users who publish content (Videos, Immersive Media, Podcasts, etc) on the platform under the VideoWiki platform. Their quality contributions are compensated with VWX tokens. Creator will be referred to the primary author most commonly or for equally shared ownership between multiple authors. If the share in unequal then primary author will be the creator and other parties referred as collaborators.
2. **Collaborators/Co-Authors**: Users who participate to author the content by creating the master version collaboratively. These can come in at the start of the project as Co-Authors or later accepted by the creator with permissions to update master branch and get credits as partial owners of the original content and are editors on the master branch.
3. **Contributors**: Users who add contributions to a previously authored content by forking it and creating versions. These are different than the collaborators only in the manner that they come in as collaborators after the first version was released to publishing. Contributors come by adding more information to an already created piece of media published on platform.
4. **Curators**: These are trusted guides for different subject areas on VideoWiki. Curators discover and share high-quality content within their subject areas (branches) and tags and make the content easy to discover for other users. Anyone can be part of Curator circle that is curating one subject area. Curators can curate curriculum style playlists of related media and collect similar media together under specific tags for better visibility.
5. **Critics**: Individuals who would provide editorial support to community in exchange for VWX tokens. They flag the media for errors and put to sandbox for community review.
6. **Coders**: Decentralized app developers who want to contribute to the platform or ecosystem. As the platform grows, our team will open up a developer platform that will allow engineers to build third-party applications, integrations, and features for VideoWiki.
7. **Committee**: A committee of the project governance rests responsible for the early phases of development, role assignments and running the project to take off to a later DAO architecture that comprises of initial founders, thought leaders and pioneers who put effort in building the initial organizational infrastructures.
8. **Consumers**: The general users of the platform who are the consuming the content created on it, these are students, learners or anyone who is not a creator yet. These consumers enjoy the immersive content created by users and contribute to the ecosystem by viewing, commenting, sharing and tipping creators.
9. **Companies**: Media Publications can partner with VideoWiki to set up independent publishing, individual prices for content, create a sustainable business model, and build reputation with our users.

Consensus based ethos
We follow the established and time-tested mechanisms with Wikipedia guidelines on fair treatment of editorials and revisions. Taking that as a foundation, our community has a very strong buy-in to consensus decision-making, underscored by guidelines such as Consensus, and Voting is not a substitute for discussion. **Consensus** is not the same as **majority**, it signifies that the concerns and views of minorities should be considered in the attempt to gain a decision which reflects community values and which most can live by to some extent or other. Most policies and procedures also develop and become refined in this same manner.

The time taken to reach some decisions is often considered to be outweighed by the wide agreement when decisions are reached. Critically, article by article, our editing ethos will strongly encourage the incorporation of views in a policy-compliant encyclopaedic style, when they meet content criteria, and the seeking of independent others' input when consensus is unclear. Even in the event of dispute and escalation, the process remains the same -- even Arbitration Committee decisions are based upon communal input, consensus, and transparency.

8. Use Cases

"VideoWiki" is the educational use-case being strongly supported and built on the protocol. However, the tech "**as an enabler**" can solve many other challenges. These are open use-cases that can be built as applications on top of our technology. We invite developers to build solutions using this protocol and provide grants and bounties to come up with working solutions.

- For Students and learners
- For Lecturers/Instructors
- For Researchers
- For Journalists and Media professionals
- For Corporate L&D
- For Learning Disabilities
- For Process Guidebooks and SOPs
- For Governmental record keeping
- For Learning Disabilities
- For Process Guidebooks and SOPs
- For Employee Communication and Collaboration
- For Sales and Business Development
- For Marketing Teams
- For Customer Experience Teams
- For Product Development Teams

For Students and learners

- Create videos for social networks (describe your feelings and thoughts and get an engaging video, add notes and vlog), create and share video messages.
- Create video presentations, try changing written text answers for video storytelling.
- Analytical text reports are ineffective in the age of visuals.
- Create an effective analytic video (visual comparison, video charts).
- Create and share notes: Notes can be recorded, shared and published.

For Lecturers / Instructors

- Open your Pandora's box with VideoWiki:
- Teaching and Assignments: VideoWiki is a real Learning Autopilot to create an immersive training video. Generate students' imaginations and shape mindset videos for creativity.
- Classroom Lectures: A video lecture is not a video message to students, but an immersive system, where video is a way of forming knowledge through practical informational influence.
- Exams: Theoretical tests are far from practice, create video tests, form the vision of objects of the professional environment.
- Share your knowledge: Cultivate academic volunteering for the well-being of the world, share your knowledge, and create unique educational video content and make money on it.
- Notes sharing: We know the technology of brainstorming, brainwriting, the era of video driven decision making has come.

For Researchers

- Open science and open innovation: The use of research videos of preparation, conduct, and experimental results. So many researchers are working in parallel on one problem: someone finds the answer faster, writes a scientific article, waits for the review process, publication, while others see this solution. But this aspect is so necessary for other researchers to solve other problems. They wait, sometimes without waiting. Science can become more dynamic if you use research videos (by the way, they are much easier to protect than scientific text, which lends itself to rewriting and copyright hundreds of times).
- Digital Collaboration: Collaborative paper writing is now turned to collaborative video research.
- Video Logs: More immersive and impressive video note taking and Vlogging of the progress. Science fiction movies like Star Trek and Avatar style record keeping is now here.

For Journalists and Media professionals

- News information is becoming outdated by the minute. The enormous work of journalists is so short-lived, but the video clips created can be used to create historical reviews, geographical, political studies. After all, seeing once is better than hearing a hundred times.
- Fake News Detection and Ownership: Fighting fake news by enabling open communities, in turn building a process of compelling readers to recalibrate their own opinions – to look at the world dynamically and flexibly.
- Censor-less Protocol: Information is key and one who controls information controls the world. Using VideoWiki decentralized protocol you can build applications that avoid a centralized censor authority rather make the information free and unbiased to any view.

For Corporate L&D

- Training Materials for Micro-Learning: Micro sized videos that can engage the attention for a short time and give periodic bursts of learning to users. Easy creation and rapid distribution.
- Rapid creation and re-edits: Courses can be updated on the fly each time the content is to be revised, making relevant updates to avoid your content to go obsolete for corporate induction, HR SOPs and multiple other regular processes.
- Portfolios: Video resume, video business cards, video presentations of projects are a new business culture of video communication and aesthetic brand formation.
- Security: Encryption of data using video objects.

For Learning Disabilities

- Accessibility Videos: Solutions for Distracted Attention Syndrome, Dyslexia, Alzheimer's Syndrome, and others that require rapid, participatory intervention to build understanding and perception. Video review is an opportunity to restore the cognitive, informational gap caused by a lesion of the central nervous system or a part of the brain.
- Engaged Learning with Video aids: Creating apps and media which is focused on accessibility and increases cognitive engagement of a learner.

For Process Guidebooks and SOPs

- The guidebook gives way to the mentor TED talks. Search for motivational speeches by keywords and create energized practical guides.

For Employee Communication and Collaboration

- The meeting recaps: Quickly recap meetings and collaborative brainstorms to share key takeaways with others and help with idea retention.
- Team knowledge sharing: Capture new lessons learned and best practices in a simple format that others can learn from, better remember, and revisit as needed.
- Executive updates: Use video to keep your teams engaged and up-to-date with company updates from leadership.
- Sales enablement: Record a quick walkthrough for your sales team on the specific resources that are available to support them and their conversations with customers.
- Virtual introductions: send a one-to-one personal message introducing yourself to a new colleague.
- New hire introductions: put a face to a name and send a quick video introducing yourself to multiple teams across your company.

For Sales and Business Development

- A warmer cold outreach: Use webcam or screen capture videos as a better way to reach out to cold prospects and to explain exactly how you can help their business.
- Account executive introduction: introduce yourself to new accounts with a short video.
- Pre-meeting connect: Reduce meeting cancellation rates by sending a video message before your meeting to confirm the details and build a more human connection.
- Post-meeting recap: Following a customer phone call or meeting, record a short video to summarize key points and next steps that they can easily forward to other colleagues.
- Email signature video: Up your email signature game by recording a short video to introduce yourself and your role. Add it to your email signature with a hyperlinked thumbnail image.
- Event follow-up: reach out to new leads following your events with a short video.

For Marketing Teams

- Campaign review: Use video versus a long email or presentation to share key campaign results with your team or leadership.
- Content marketing: Before you launch a new content asset, record a video to educate your internal teams on why it’s important and how they can help to promote it.
- Video newsletter: Spice up your newsletter by recording and embedding custom videos to complement your email communications.

- Digital marketing: Use screen capture videos to share website updates with internal teams in a way that enables you to explain both the what and the why.
- Event marketing: Use the power of video to improve outreach to potential partners, sponsors, speakers, and attendees for your upcoming events.
- Content promotion: Embed a promo video in campaign emails or landing pages to provide a high-level overview of your content.
- Creative design: Create short videos to solicit or provide feedback on product design, website design, and other creative assets.
- Press pitches: Stand out with reporters by pitching new story ideas in a way that introduces yourself and creates a more personal connection.

For Customer Experience Teams

- Faster, better customer support: Stop sending text-based instructions and screen shots. Instead, start using video to show exactly how to solve an issue in an easy-to-follow format.
- Welcome new customers: Put a face to your name by welcoming new customers with a personal video message from their account manager, support rep, or your execs.
- Customer onboarding: Improve the customer experience by outlining the onboarding process for your new customers.
- Ongoing customer communications: Update your customers on company news and product updates with video messages that stand out and get noticed.
- Congratulate customers: Share a video to celebrate a customer's special moment or milestone.

For Product Development Teams

- Project management: Share video messages to clearly summarize project plans and to keep teammates aligned on key deliverables.
- Design and UX reviews: Use custom screen capture videos to solicit or provide feedback on user experience design updates.
- Product feedback solicitation: Encourage other teams and customers to provide product feedback via custom videos that clearly show, and explain, their perspective.

9. The Market (Product-Market Fit)

- Size Target
- Market Potential
- Impact of COVID-19
- Content, Storage and Retrieval Markets

10. VideoWiki Incentivisation Model (Business Economics)

- The Contribution Economy
- User Rewards
- Content Development Rewards

11. Competitive Landscape

12. Strategy Roadmap